STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _____ 10/534,130 _____
Source: _____ PCT _____
Date Processed by STIC: _____ 01/31/2006 _____

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.2.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
     http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio **(<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):**
    U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

Raw Sequence Listing Error Summary

ERROR DETECTED   SUGGESTED CORRECTION   SERIAL NUMBER: _10/534,130_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1____Wrapped Nucleics   The number/text at the end of each line "wrapped" down to the next line. This may occur if your file
     Wrapped Aminos   was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will
                        prevent "wrapping."

2____Invalid Line Length The rules require that a line **not exceed 72 characters** in length. This includes white spaces.

3____Misaligned Amino   The numbering under each 5th amino acid is misaligned. Do **not** use tab codes between numbers;
     Numbering   use **space characters**, instead.

4____Non-ASCII   The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please
                        ensure your subsequent submission is saved in ASCII text.**

5____Variable Length   Sequence(s)____contain n's or Xaa's representing more than one residue. **Per Sequence Rules,
                        each n or Xaa can only represent a single residue.** Please present the **maximum** number of each
                        residue having variable length and indicate in the <220>-<223> section that some may be missing.

6____PatentIn 2.0   A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
     "bug"   sequences(s) _____. Normally, PatentIn would automatically generate this section from the
                        previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to
                        the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for
                        Artificial or Unknown sequences.**

7____Skipped Sequences Sequence(s) _____ missing. If intentional, please insert the following lines for **each** skipped sequence:
     (OLD RULES)   (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
                        (i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
                        (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
                        This sequence is intentionally skipped

                        Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8____Skipped Sequences Sequence(s) _____ missing. If **intentional**, please insert the following lines for **each** skipped sequence.
     (NEW RULES)   <210> sequence id number
                        <400> sequence id number
                        000

9____Use of n's or Xaa's Use of n's and/or Xaa's have been detected in the Sequence Listing.
     (NEW RULES)   Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
                        In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____Invalid <213>   Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or
      Response   scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or
                        is Artificial Sequence

11____Use of <220>   Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
                        Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or
                        "Unknown." Please explain source of genetic material in <220> to <223> section.
                        (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12____PatentIn 2.0   Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file,
      "bug"   resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence
                        listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13____Misuse of n/Xaa   "n" can **only** represent a single nucleotide; "Xaa" can **only** represent a single amino acid

AMC – Biotechnology Systems Branch – 09/09/2003

PCT

RAW SEQUENCE LISTING   DATE: 01/31/2006
PATENT APPLICATION: US/10/534,130   TIME: 16:12:34

              Input Set : **A:\PTO.TS.31.txt**
              Output Set: **N:\CRF4\01312006\J534130.raw**

```
 3 <110> APPLICANT: Ian Hector Frazer
 5 <120> TITLE OF INVENTION: A method for optimising gene expression using
 6       synonymous codon optimisation
 8 <130> FILE REFERENCE: 21415-0015US
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/534,130
11 <141> CURRENT FILING DATE: 2005-05-06
13 <150> PRIOR APPLICATION NUMBER: PCT/AU2003/001487
14 <151> PRIOR FILING DATE: 2003-11-10
16 <150> PRIOR APPLICATION NUMBER: USSN 60/425,163
17 <151> PRIOR FILING DATE: 2002-11-08
19 <160> NUMBER OF SEQ ID NOS: 126·
21 <170> SOFTWARE: PatentIn version 3.2
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 714
25 <212> TYPE: DNA
26 <213> ORGANISM: Artificial Sequence
28 <220> FEATURE:
29 <223> OTHER INFORMATION: Humanised GFP
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (1)..(711)
36 <400> SEQUENCE: 1
37 agc aag ggc gag gaa ctg ttc act ggc gtg gtc cca att ctc gtg gaa      48
38 Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu
39 1               5                   10                  15
41 ctg gat ggc gat gtg aat ggg cac aaa ttt tct gtc agc gga gag ggt      96
42 Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly
43               20                  25                  30
45 gaa ggt gat gcc aca tac gga aag ctc acc ctg aaa ttc atc tgc acc     144
46 Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr
47               35                  40                  45
49 act gga aag ctc cct gtg cca tgg cca aca ctg gtc act acc ttc tct     192
50 Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe Ser
51       50                  55                  60
53 tat ggc gtg cag tgc ttt tcc aga tac cca gac cat atg aag cag cat     240
54 Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His
55 65                  70                  75                  80
57 gac ttt ttc aag agc gcc atg ccc gag ggc tat gtg cag gag aga acc     288
58 Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr
59               85                  90                  95
61 atc ttt ttc aaa gat gac ggg aac tac aag acc cgc gct gaa gtc aag     336
62 Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys
63               100                 105                 110
```

Does Not Comply
Corrected Diskette Needed
(pg-3,4,5)

```
65 ttc gaa ggt gac acc ctg gtg aat aga atc gag ctg aag ggc att gac     384
66 Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp
67         115                 120                 125
69 ttt aag gag gat gga aac att ctc ggc cac aag ctg gaa tac aac tat     432
70 Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr
71     130                 135                 140
73 aac tcc cac aat gtg tac atc atg gcc gac aag caa aag aat ggc atc     480
74 Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile
75 145                 150                 155                 160
77 aag gtc aac ttc aag atc aga cac aac att gag gat gga tcc gtg cag     528
78 Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln
79   .             165                 170                 175
81 ctg gcc gac cat tat caa cag aac act cca atc ggc gac ggc cct gtg     576
82 Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val
83             180                 185                 190
85 ctc ctc cca gac aac cat tac ctg tcc acc cag tct gcc ctg tct aaa     624
86 Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys
87         195                 200             .   205
89 gat ccc aac gaa aag aga gac cac atg gtc ctg ctg gag ttt gtg acc     672
90 Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr
91     210                 215                 220           . .
93 gct gct ggg atc aca cat ggc atg gac gag ctg tac aag tga            714
94 Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
95 225                 230                 235
98 <210> SEQ ID NO: 2
99 <211> LENGTH: 237
100 <212> TYPE: PRT
101 <213> ORGANISM: Artificial Sequence
103 <220> FEATURE:
104 <223> OTHER INFORMATION: Humanised GFP
106 <400> SEQUENCE: 2
108 Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu
109 1               5                   10                  15
112 Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly
113             20                  25                  30
116 Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr
117         35                  40                  45
120 Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe Ser
121     50                  55                  60
124 Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His
125 65                  70                  75                  80
128 Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr
129             85                  90                  95
132 Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys
133             100                 105                 110
136 Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp
137         115                 120                 125
140 Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr
141     130                 135                 140
```

RAW SEQUENCE LISTING   DATE: 01/31/2006
PATENT APPLICATION: US/10/534,130   TIME: 16:12:34

Input Set : A:\PTO.TS.31.txt
Output Set: N:\CRF4\01312006\J534130.raw

```
144 Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile
145 145             150             155             160
148 Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln
149·            165             170             175
152 Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val
153         180             185             190
156 Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys
157     195             200             205
160 Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr
161     210             215             220
164 Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
165 225             230             235
168 <210> SEQ ID NO: 3
169 <211> LENGTH: 18
170 <212> TYPE: DNA
171 <213> ORGANISM: Artificial Sequence
173 <220> FEATURE:
174 <223> OTHER INFORMATION Ala(GCA)x6
177 <220> FEATURE:
178 <221> NAME/KEY: CDS
179 <222> LOCATION: (1)..(18)
181 <400> SEQUENCE: 3
182 gca gca gca gca gca gca                                       18
183 Ala Ala Ala Ala Ala Ala
184 1             5
187 <210> SEQ ID NO: 4
188 <211> LENGTH: 6
189 <212> TYPE: PRT
190 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: Ala(GCA)x6
195 <400> SEQUENCE: 4
197 Ala Ala Ala Ala Ala Ala
198 1             5
201 <210> SEQ ID NO: 5
202 <211> LENGTH: 18
203 <212> TYPE: DNA
204 <213> ORGANISM: Artificial Sequence
206 <220> FEATURE:
207 <223> OTHER INFORMATION: Ala(GCG)x6
210 <220> FEATURE:
211 <221> NAME/KEY: CDS
212 <222> LOCATION: (1)..(18)
214 <400> SEQUENCE: 5
215 gcg gcg gcg gcg gcg gcg                                       18
216 Ala Ala Ala Ala Ala Ala
217 1             5
220 <210> SEQ ID NO: 6
221 <211> LENGTH: 6
```

[handwritten annotation]: → Invalid Response. Pls Explain the Source of genetic Material. See Item #11 on Error Summary Sheet.

[handwritten annotation]: → Same Error

Input Set : **A:\PTO.TS.31.txt**
Output Set: **N:\CRF4\01312006\J534130.raw**

```
222 <212> TYPE: PRT
223 <213> ORGANISM: Artificial Sequence
225 <220> FEATURE:
226 <223> OTHER INFORMATION: Ala(GCG)x6
228 <400> SEQUENCE: 6
230 Ala Ala Ala Ala Ala Ala
231 1               5
234 <210> SEQ ID NO: 7
235 <211> LENGTH: 18
236 <212> TYPE: DNA
237 <213> ORGANISM: Artificial Sequence
239 <220> FEATURE:
240 <223> OTHER INFORMATION: Ala(GCT)x6
243 <220> FEATURE:
244 <221> NAME/KEY: CDS
245 <222> LOCATION: (1)..(18)
247 <400> SEQUENCE: 7
248 gct gct gct gct gct gct                                      18
249 Ala Ala Ala Ala Ala Ala
250 1               5
253 <210> SEQ ID NO: 8
254 <211> LENGTH: 6
255 <212> TYPE: PRT
256 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Ala(GCT)x6
261 <400> SEQUENCE: 8
263 Ala Ala Ala Ala Ala Ala
264 1               5
267 <210> SEQ ID NO: 9
268 <211> LENGTH: 18
269 <212> TYPE: DNA
270 <213> ORGANISM: Artificial Sequence
272 <220> FEATURE:
273 <223> OTHER INFORMATION: Ala(GCC)x6
276 <220> FEATURE:
277 <221> NAME/KEY: CDS
278 <222> LOCATION: (1)..(18)
280 <400> SEQUENCE: 9
281 gcc gcc gcc gcc gcc gcc                                      18
282 Ala Ala Ala Ala Ala Ala
283 1               5
286 <210> SEQ ID NO: 10
287 <211> LENGTH: 6
288 <212> TYPE: PRT
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: Ala(GCC)x6
294 <400> SEQUENCE: 10
```

Same error

RAW SEQUENCE LISTING   DATE: 01/31/2006
PATENT APPLICATION: US/10/534,130   TIME: 16:12:34

Input Set : A:\PTO.TS.31.txt
Output Set: N:\CRF4\01312006\J534130.raw

```
296 Ala Ala Ala Ala Ala Ala
297 1                    5
300 <210> SEQ ID NO: 11
301 <211> LENGTH: 18
302 <212> TYPE: DNA
303 <213> ORGANISM: Artificial Sequence
305 <220> FEATURE:
306 <223> OTHER INFORMATION: Arg(AGA)x6
309 <220> FEATURE:
310 <221> NAME/KEY: CDS
311 <222> LOCATION: (1)..(18)
313 <400> SEQUENCE: 11
314 aga aga aga aga aga aga                                    18
315 Arg Arg Arg Arg Arg Arg
316 1                    5
319 <210> SEQ ID NO: 12
320 <211> LENGTH: 6
321 <212> TYPE: PRT
322 <213> ORGANISM: Artificial Sequence
324 <220> FEATURE:
325 <223> OTHER INFORMATION: Arg(AGA)x6
327 <400> SEQUENCE: 12
329 Arg Arg Arg Arg Arg Arg
330 1                    5
333 <210> SEQ ID NO: 13
334 <211> LENGTH: 18
335 <212> TYPE: DNA
336 <213> ORGANISM: Artificial Sequence
338 <220> FEATURE:
339 <223> OTHER INFORMATION: Arg(CGA)x6
342 <220> FEATURE:
343 <221> NAME/KEY: CDS
344 <222> LOCATION: (1)..(18)
346 <400> SEQUENCE: 13
347 cga cga cga cga cga cga                                    18
348 Arg Arg Arg Arg Arg Arg
349 1                    5
352 <210> SEQ ID NO: 14
353 <211> LENGTH: 6
354 <212> TYPE: PRT
355 <213> ORGANISM: Artificial Sequence
357 <220> FEATURE:
358 <223> OTHER INFORMATION: Arg(CGA)x6
360 <400> SEQUENCE: 14
362 Arg Arg Arg Arg Arg Arg
363 1                    5
366 <210> SEQ ID NO: 15
367 <211> LENGTH: 18
368 <212> TYPE: DNA
```

Same Error

VERIFICATION SUMMARY   DATE: 01/31/2006
PATENT APPLICATION: **US/10/534,130**   TIME: 16:12:35

Input Set : **A:\PTO.TS.31.txt**
Output Set: **N:\CRF4\01312006\J534130.raw**

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number